

FATEMEH NARGESIAN

University of Rochester, Department of Computer Science

✉ fnargesian@rochester.edu

<https://fnargesian.com/>

RESEARCH INTERESTS

The first theme in my research revolves around *data acquisition for AI*. I design algorithms and build systems that efficiently discover relevant data in lakes of heterogeneous data sources and select and integrate sources into a dataset that satisfies the schema and distribution requirements for accurate and fair AI. The second theme is *scientific time-series analysis*, where I aim to develop ways of scaling mathematical models for time-series correlation to large-scale and fast-updating scientific time-series to enable interactive data exploration in domains such as climate science and neuroscience.

ACADEMIC APPOINTMENT

University of Rochester, Department of Computer Science

Assistant Professor, July 2019 - present

EDUCATION

University of Toronto

PhD, Computer Science, 2019

Advisor: Renée J. Miller

MSc, Computer Science, University of Ottawa, Canada

MSc, Artificial Intelligence, Sharif University of Technology, Iran

BSc, Computer Engineering, Shahid Beheshti University, Iran

INDUSTRY RESEARCH EXPERIENCE

IBM T. J. Watson Research Center, NY, USA

Research Intern at Automated Machine Learning Group, 2016

IBM T. J. Watson Research Center, NY, USA

Research Intern at Stream Analytics Group, 2014

McGill University, QC, Canada

Researcher at Clinical Informatics Research Group, 2011

PUBLICATIONS

Sampling over Union of Joins.

Y. Liu, Y. Xu, F. Nargesian. CoRR abs/2303.00940, 2023.

KOIOS: Top-k Semantic Overlap Set Search.

P. Mundra, J. Zhang, F. Nargesian, N. Augsten. ICDE, 2023.

Matching Roles from Temporal Data.

L. Bornemann, T. Bleifuß, D. V. Kalashnikov, F. Nargesian, F. Naumann, D. Srivastava. SIGMOD, 2023.

Responsible Data Integration: Next-generation Challenges. (Tutorial)

F. Nargesian, A. Asudeh, H. V. Jagadish. WSDM 2023.

TSUPY: Dynamic Climate Network Analysis Library.

J. Liu, Y. Xu, F. Nargesian, G. Ghoshal. CIKM 2022: 4926-4930.

Distribution-aware Data Acquisition in Online Data Markets.

A. Asudeh, F. Nargesian. PVLDB 15(11): 3137-3144 (2022).

TSUBASA: Climate Network Construction on Historical and Real-Time Data.

Y. Xu, J. Liu, F. Nargesian. SIGMOD 2022: 286-295.

Responsible Data Integration: Next-generation Challenges. (Tutorial)

F. Nargesian, A. Asudeh, H. V. Jagadish. SIGMOD 2022: 2458-2464.

Tailoring Data Source Distributions for Fairness-aware Data Integration.

F. Nargesian, A. Asudeh, H. V. Jagadish. PVLDB 14(11): 2519-2532 (2021).

RONIN: Data Lake Exploration.

P. Ouellette, A. Sciortino, F. Nargesian, B. Ghadiri-Bashardoost, E. Zhu, K. Pu, R. J. Miller. PVLDB 14(12): 2863-2866 (2021).

AWLCO: All-Window Length Co-Occurrence.

J. Sobel, N. Bertram, C. Ding, F. Nargesian, D. Gildea. Symposium on Combinatorial Pattern Matching, 24:1-24:21, 2021.

Data Lake Organization.

F. Nargesian, K. Pu, B. Ghadiri-Bahshardoost, E. Zhu, R. J. Miller. TKDE, 2021.

Organizing Data Lakes for Navigation.

F. Nargesian, K. Q. Pu, E. Zhu, B. G. Bashardoost, R. J. Miller. SIGMOD 2020: 1939-1950.

Knowledge Translation.

B. Ghadiri-Bashardoost, R. J. Miller, K. Lyons, F. Nargesian. PVLDB 13(11), 2020.

Loki: Streamlining Integration and Enrichment.

W. Spoth, P. Kumari, O. Kennedy, F. Nargesian. HILDA@SIGMOD, 2020.

Data Lake Management: Challenges and Opportunities. (Tutorial)

F. Nargesian, E. Zhu, R. J. Miller, Ken Q. Pu. PVLDB 12(12): 1986-1989 (2019).

JOSIE: Overlap Set Similarity Search for Finding Joinable Tables in Data Lakes.

E. Zhu, D. Deng, F. Nargesian, R. J. Miller. SIGMOD 2019: 847-864.

Making Open Data Transparent: Data Discovery on Open Data.

R. J. Miller, F. Nargesian, E. Zhu, Christina Christodoulakis, K. Q. Pu, Periklis Andritsos. IEEE Data Engineering Bulletin, 41(2): 59-70 (2018).

Table Union Search on Open Data.

F. Nargesian, E. Zhu, K. Q. Pu, R. J. Miller. PVLDB 11(7), 2018.

Dataset Evolver: An Interactive Feature Engineering Notebook. (Demo)

F. Nargesian, U. Khurana, H. Samulowitz, D. S. Turaga, T. Pedapati. AAAI 2018: 8212-8213.

Interactive Navigation of Open Data Linkages. (*Best Demo Award*)

E. Zhu, K. Q. Pu, F. Nargesian, R. J. Miller. PVLDB 10(12): 1837-1840 (2017).

Learning Feature Engineering for Classification.

F. Nargesian, H. Samulowitz, U. Khurana, E. B. Khalil, D. S. Turaga. IJCAI 2017: 2529-2535.

Automating Feature Engineering.

U. Khurana, F. Nargesian, H. Samulowitz, D. S. Turaga, E. B. Khalil. AI for Data Science Workshop, NeurIPS, 2016.

LSH Ensemble: Internet-Scale Domain Search.

E. Zhu, F. Nargesian, K. Q. Pu, R. J. Miller. PVLDB 9(12), 2016.

LinkedCT Live: Platform for Online Curation of Clinical Trials Data.

O. Hassanzadeh, R. J. Miller, F. Nargesian, E. Zhu. ISWC, 2015.

SOFIA: An Analytics Recommendation System.

F. Nargesian, A. Biem, P. Jain, S. Parthasarathy, D. S. Turaga. ISWC, 2015.

Data-driven recommendations for exploratory query formulation.

F. Nargesian, SIGMOD PhD Symposium 2014: 31-35.

Managing and Mapping Data Lineage for Business Intelligence and Analytics Applications in Health Care.

M. Azarm, F. Nargesian, L. Peyton, i-Society: 120-126, 2011.

Tool Support and Data Management for Business Analytics Applications in Healthcare.

M. Azarm, F. Nargesian, L. Peyton, International Journal for Infonomics (IJI) 4(4): 484-493, 2011.

TEACHING EXPERIENCE

University of Rochester

Topics in Data Management (CSC 576/578): Fall 2019-2020.

Data Systems Management (CSC/DSC 263/463): Spring 2020-2023.

Undergraduate Problem Seminar (CSC200/200H): Spring 2022-2023.

University of Toronto

Introduction to Databases, Fall 2015 (Sessional Instructor).

AWARDS

NSF IIS-2107050 (2021 - 2025): III : Medium: Collaborative Research: From Open Data to Open Data Curation. Fatemeh Nargesian (UR), Renée J. Miller (Northeastern), and Boris Glavic (IIT). \$1.2M (UR portion: \$344,874K).

Goergen Institute for Data Science Seed Funding Program (2021-2022): Interactive Climate Network Exploration over Real-Time Data. Fatemeh Nargesian (Department of Computer Science, UR) and Gourab Ghoshal (Department of Physics and Astronomy, UR). \$20K

Wadsworth C. Sykes Engineering Faculty Award for designing "Climate Data Management" course. \$2.5K

Best Demo Award (2017) - 43th International Conference on Very Large Data Bases (VLDB)

International Francophone scholarship (2008-2011) - University of Ottawa

ACADEMIC SERVICES

Program Committee Member

VLDB 2022-2024, SIGMOD 2022-2024, SIGMOD Reproducibility 2023, ICDE 2022, EDBT Demo 2022, VLDB Reproducibility 2022, VLDB PhD Workshop 2020-2021, CIKM 2020, Scalable Data Integration round-table at VLDB 2021.

Journal Reviewer

TKDE 2019-2021, VLDBJ 2022.

Conference Organizer

Registration Chair at VLDB 2023, Undergraduate Research Co-chair at SIGMOD 2023.

Department Committee Member at University of Rochester

Colloquium committee 2020-2021, admission committee 2019-2022, diversity committee 2019-2022, teaching faculty recruiting committee 2022.

Grant Reviewer

NSF CISE SMALL 2021 (panelist), MITACS 2021 (reviewer)

PATENTS

USPTO Application #82017030001: Methods and Systems for Feature Engineering